C L A I M S

1. A sentence classification device
characterized by comprising:

a term list having a plurality of terms each
comprising not less than one word;

DT matrix generation means for generating a DT
matrix two-dimensionally expressing a relationship
between each document contained in a document set and
said each term;

DT matrix transformation means for generating
a transformed DT matrix having clusters having blocks of
associated documents by transforming the DT matrix
obtained by said DT matrix generation means on the basis
of a DM decomposition method used in a graph theory; and

classification generation means for generating
classifications associated with the document set on the
basis of a relationship between each cluster on the
transformed DT matrix obtained by said DT matrix
transformation means and said each document classified
according to the clusters.

2. A sentence classification device according
to claim 1, characterized in that said classification
generation means comprises document classification means
for outputting, for each cluster on the transformed DT
matrix obtained by said DT matrix transformation means,
documents belonging to the cluster as the same
classification.

3. A sentence classification device according to claim 2, characterized by further comprising label generation means for outputting each term strongly connected to each document belonging to said arbitrary cluster as a label indicating a classification of the cluster.

4. A sentence classification device according to claim 2, characterized by further comprising document organization means for sequentially outputting documents belonging to said arbitrary cluster or all documents in an arrangement order of the documents in the transformed DT matrix.

5. A sentence classification device according to claim 2, characterized by further comprising summary generation means for outputting, as a summary of said arbitrary document, a sentence of sentences constituting the document which contains a term strongly connected to the document.

6. A sentence classification device according to claim 2, characterized by further comprising:

term list edition means for adding or deleting an arbitrary term with respect to the term list; and

index generation means for making said DT matrix generation means generate DT matrices by using term lists before and after edition by said term list edition means, and generating and outputting an index indicating validity of the edition from the DT matrices.

7. A sentence classification device according
2 to claim 1, characterized in that said classification
3 generation means comprises:

4 virtual representative document generation
5 means for generating a virtual representative document,
6 for each cluster on a transformed DT matrix, from a term
7 of each document belonging to the cluster; and

8 large classification generation means for
9 generating a large classification of documents by
10 repeatedly performing clustering processing of setting a
11 DT matrix generated by said DT matrix generation means
12 in an initial state, causing said virtual representative
13 document generation means to generate a virtual
14 representative document for each cluster on a
15 transformed DT matrix generated from the DT matrix by
16 said DT matrix transformation means, generating a new DT
17 matrix used for next clustering processing by adding the
18 virtual representative document to the transformed DT
19 matrix and deleting documents belonging to the cluster
20 of the virtual representative document from the
21 transformed DT matrix, and outputting, for said each
22 cluster, information associated with the documents
23 constituting the cluster as large classification data.

8. A sentence classification device according
2 to claim 7, characterized in that said large
3 classification generation means terminates repetition of
4 the clustering processing when no cluster is obtained

5   from the transformed DT matrix in the clustering

6   processing.

         9.   A sentence classification device according

2   to claim 7, characterized by further comprising large

3   classification label generation means for, if a virtual

4   representative document is contained in a given cluster

5   of clusters obtained by the clustering processing,

6   generating a label of the cluster on which the virtual

7   representative document is based from a term strongly

8   connected to the virtual representative document.

         10.   A sentence classification method

2   characterized by comprising:

3   the DT matrix generation step of generating a

4   DT matrix two-dimensionally expressing a relationship

5   between each document contained in a document set and

6   each term of a term list having a plurality of terms

7   each comprising not less than one word;

8   the DT matrix transformation step of

9   generating a transformed DT matrix having clusters

10   having blocks of associated documents by transforming

11   the DT matrix on the basis of a DM decomposition method

12   used in a graph theory; and

13   the classification generation step of

14   generating classifications associated with the document

15   set on the basis of a relationship between each cluster

16   on the transformed DT matrix and said each document

17   classified according to the clusters.

11.   A sentence classification method according
to claim 10, characterized in that the classification
generation step comprises the document classification
step of outputting, for each cluster on the transformed
DT matrix, documents belonging to the cluster as the
same classification.

12.   A sentence classification method according
to claim 11, characterized by further comprising the
step of outputting each term strongly connected to each
document belonging to said arbitrary cluster as a label
indicating a classification of the cluster.

13.   A sentence classification method according
to claim 11, characterized by further comprising the
step of sequentially outputting documents belonging to
said arbitrary cluster or all documents in an
arrangement order of the documents in the transformed DT
matrix.

14.   A sentence classification method according
to claim 11, characterized by further comprising the
step of outputting, as a summary of a document, a
sentence of sentences constituting said arbitrary
document which contains a term strongly connected to the
document.

15.   A sentence classification method according
to claim 11, characterized by further comprising:
        the step of adding or deleting an arbitrary
term with respect to the term list; and

5   the step of generating DT matrices by using

6   term lists before and after edition, and generating and

7   outputting an index indicating validity of the edition

8   from the DT matrices.

16.   A sentence classification method according

2   to claim 10, characterized in that the classification

3   generation step comprises:

4   the virtual representative document generation

5   step of generating a virtual representative document,

6   for each cluster on a transformed DT matrix, from a term

7   of each document belonging to the cluster; and

8   the large classification generation step of

9   generating a large classification of documents by

10 repeatedly performing clustering processing comprising

11 the step of setting a DT matrix generated in the DT

12 matrix generation step in an initial state, generating a

13 virtual representative document in the virtual

14 representative document generation step for each cluster

15 on a transformed DT matrix generated from the DT matrix

16 in the DT matrix transformation step, the step of

17 generating a new DT matrix used for next clustering

18 processing by adding the virtual representative document

19 to the transformed DT matrix and deleting documents

20 belonging to the cluster of the virtual representative

21 document from the transformed DT matrix, and the step of

22 outputting, for said each cluster, information

23 associated with the documents constituting the cluster

24   as large classification data.

      17.   A sentence classification method according

2   to claim 16, characterized in that in the large

3   classification generation step, repetition of the

4   clustering processing is terminated when no cluster is

5   obtained from the transformed DT matrix in the

6   clustering processing.

      18.   A sentence classification method according

2   to claim 16, characterized by further comprising the

3   large classification label generation step of, if a

4   virtual representative document is contained in a given

5   cluster of clusters obtained by the clustering

6   processing, generating a label of the cluster on which

7   the virtual representative document is based from a term

8   strongly connected to the virtual representative

9   document.